Preprint CheMIn Technical Reports ISSN 3061-1075

Four Decades of AI in Chemistry Education

Aleksandar Kondinski [1]

released: 11 October 2026

[1] Institute of Physical and
Theoretical Chemistry
Graz University of Technology
Stremayrgasse 9, 8010 Graz, Austria

Preprint No. 9



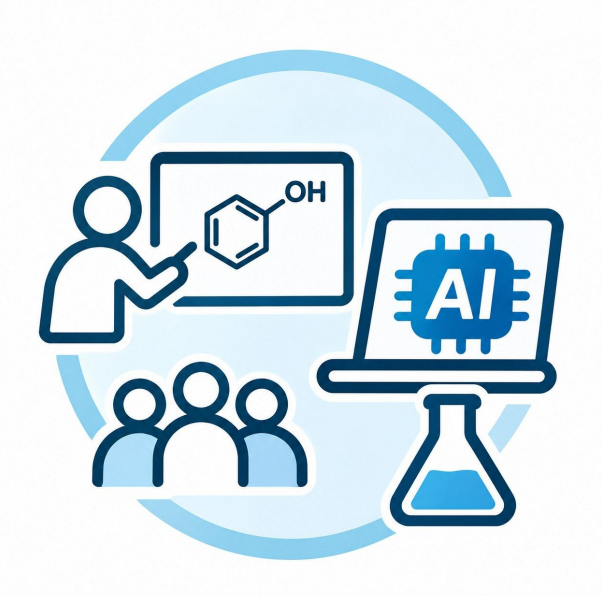
Keywords: AI in education, chemistry education, generative AI, machine learning, personalised learning

Edited by
Chemical Modelling and Informatics Group
Institute of Physical and Theoretical Chemistry
Graz University of Technology
Stremayrgasse 9, 8010 Graz
Austria (Europe)

E-Mail: kondinski@tugraz.at
World Wide Web: http://chemin.at/

Abstract

Artificial intelligence has been used in chemistry education for four decades, from early rule-based tutors to today's generative AI systems. This review examines more than one hundred papers through the TPACK and SAMR frameworks and asks how far increasing technical capability has changed chemistry teaching and learning. Across machine learning, generative AI, assessment, personalised learning and laboratory instruction, most applications remain at the levels of substitution or augmentation, while fewer substantially modify the learning activity. Longitudinal and controlled evidence remains limited, and evidence for disabled and neurodivergent learners remains sparse. Overall, technical capability has advanced further than pedagogical change. The educational value of AI will depend on whether it strengthens chemical reasoning, supports authentic scientific practice and leads to durable learning.



Highlights

- AI has supported chemistry education for four decades.
- Most applications remain substitution or augmentation on SAMR.
- Generative AI changes interaction more than learning design.
- Evidence for durable learning and for equity remains sparse.

Contents

1 Introduction **3**

2 Machine Learning and Data-Driven Learning in Chemistry Education **5**

3 Generative AI and Large Language Models in Chemistry Education **7**

3.1 Capabilities and Accuracy of LLMs in Chemistry 7

3.2 Prompt Engineering and Instructional Strategies 8

4 AI-Assisted Assessment and Feedback **9**

5 Personalised and Adaptive Learning Systems **11**

6 AI in Laboratory Instruction and Simulation **13**

6.1 Generative AI in Lab Preparation and Reporting 13

6.2 Virtual Labs and Simulations 14

6.3 Data Analysis and Visualisation in Labs 15

7 Reliability, Integrity, and Equity **15**

7.1 Reliability as an Educational Risk 16

7.2 Academic Integrity and Assessment Design 17

7.3 Equity and Access 17

8 Conclusions and Outlook **18**

References **20**

1 Introduction

Artificial intelligence (AI) is often presented as a recent development in chemistry education because generative AI has made it much more visible. Its educational use, however, reaches back around four decades. One of the earliest examples is the GEORGE program, developed in the mid 1980s as a "problem solving partner" for introductory chemistry students. Students could enter problems involving quantities such as mass, volume and moles and receive explanations in natural language [14]. Even at this early stage, AI was being used to make computer-based instruction more interactive than conventional drill and practice software. What is new, then, is not the presence of AI in chemistry education, but the scale and capability of the systems now available. This longer history shifts the question from whether AI can be used in chemistry education to what, after four decades of development, it has actually changed about how chemistry is taught and learned.

Intelligent tutoring systems extended this idea in the early 1990s by connecting models of human cognition with computer-based instruction. Constructing software to teach chemical reactivity was also seen as a way to examine pedagogy because such systems required an explicit model of how students learn [6]. AI was beginning to serve two related purposes, supporting students directly and providing a way to examine the learning process itself. A parallel development occurred in chemical engineering, where a course combining symbolic and data-driven AI to teach first-principles modelling has operated for 35 years [97]. Over time, the field broadened from symbolic systems and intelligent tutors to machine learning and analytics, data-driven learning, and more recently generative and agentic AI. These developments overlap rather than replace one another, but they show a clear expansion in what AI systems can do. Large language models can now generate explanations, solve problems, analyse information and support increasingly complex scientific tasks. The rapid pace of change has made the experiences of early adopters particularly valuable [111] and has brought questions of reliability, academic integrity, dependence on AI systems and the changing role of the instructor much closer to everyday teaching.

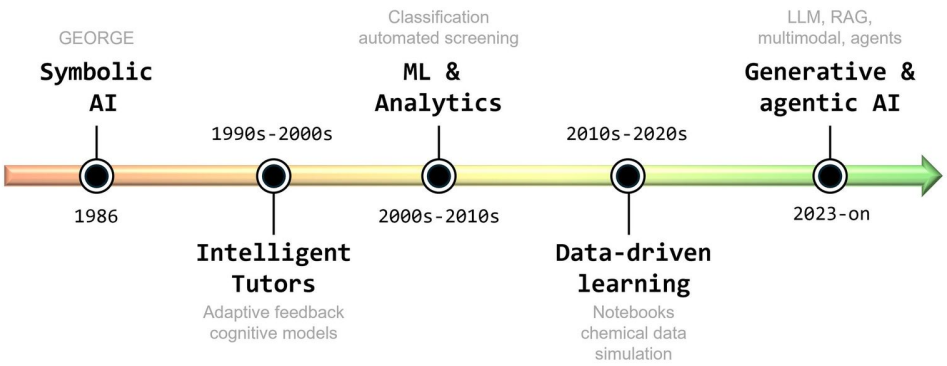


Figure 1: *Broad historical trajectory of AI in chemistry education from 1986 to 2026, highlighting shifts in emphasis from symbolic AI and intelligent tutors to machine learning and analytics, data-driven learning, and generative and agentic AI.*

These developments also matter beyond the classroom because AI is becoming part of the scientific practice students are being prepared to enter. It is already widely used in areas such as pharmacology and biochemistry for prediction and molecular design [5]. The idea of the "hybrid chemist" captures this shift by combining experimental expertise with computation, machine learning, automation and data stewardship. One proposed approach integrates these capabilities into chemistry courses through laboratory modules organised around planning, execution, measurement, updating and decision-making [9]. Against this background, this review asks how far increasing technical capability has translated into meaningful change in chemistry teaching and learning. It brings together more than one hundred papers identified across chemistry-specific and broader scholarly databases. Earlier reviews have mainly addressed more recent periods of AI adoption, broader science-education settings, or emerging educational technologies without following the longer development of AI within chemistry education [12, 21, 37]. The focus here is therefore not only on what these technologies can do, but on what they change educationally.

Across the literature, five educational roles recur: teaching and explanation, assessment and feedback, personalised and adaptive learning, laboratory instruction and simulation, and preparation for research practice. These roles cut across technological periods. Tutoring was already central to early intelligent systems [6], and automated grading appeared well before the current generative-AI wave [41]. Figure 2 summarises these roles and provides a map for much of the discussion that follows.

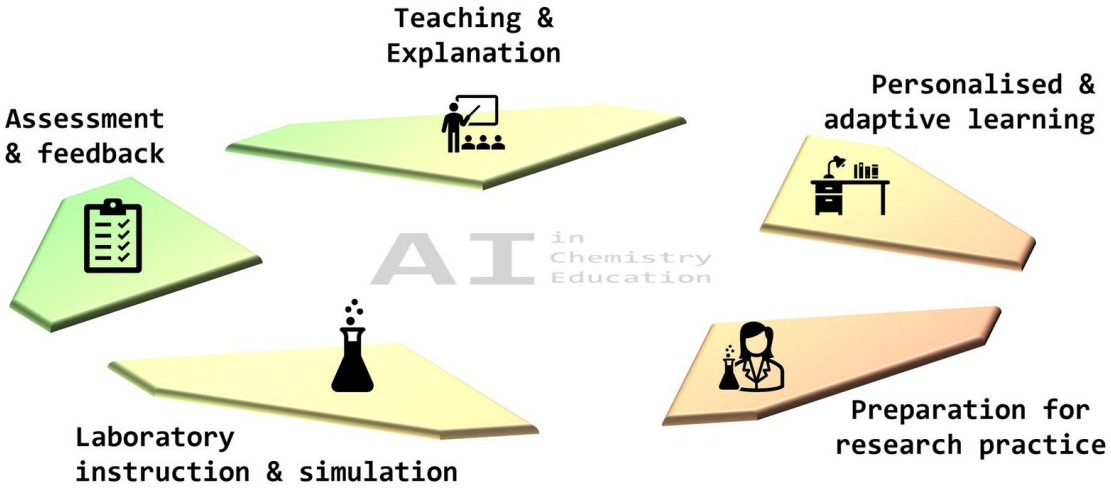


Figure 2: *Five recurring educational roles of AI across the reviewed literature.*

Two complementary frameworks guide the analysis. Technological Pedagogical Content Knowledge (TPACK) considers how technological, pedagogical and content knowledge need to come together for technology to be integrated meaningfully into teaching [62]. Here, TPACK is used mainly to consider what teachers need in order to work with AI effectively and responsibly. The Substitution, Augmentation, Modification and Redefinition (SAMR) model asks a different question. It considers how far technology changes the learning activity itself, from replacing or extending an existing task to enabling forms of learning that were previously difficult or impossible [29]. Across the literature, most

applications still support, accelerate or automate activities that already existed. Fewer substantially change how students engage with chemical knowledge, and convincing cases of genuine Redefinition remain rare. AI also introduces demands that extend beyond conventional technological knowledge, including prompting skills and awareness of bias, discrimination and hallucination [23]. SAMR levels are assigned interpretively from the learning activity described in each study rather than as a formal coding scheme. Table 1 gives representative examples.

Table 1: *Representative applications of AI in chemistry education mapped to the SAMR framework. The examples span substitution, augmentation and modification, while none clearly reaches Redefinition.*

System / study	Educational role	SAMR level	Source
Notebook-based ML workshops	Introduces ML through interactive exercises using chemical data	Augmentation	[46]
ACD/Structure Elucidator	Supports NMR structure interpretation	Augmentation	[82]
ChatGPT laboratory report drafting	Assists with laboratory report writing	Substitution	[36]
Automated text scoring	Supports scoring of open-ended responses	Augmentation	[18, 67]
AI use detection	Supports academic integrity monitoring	Augmentation	[83]
ChEdu and Professor Leodar	Course-aligned Socratic tutoring	Modification	[92, 102]
VR laboratory avatars	Provides interactive laboratory guidance	Modification	[88]
AI-generated graphic storytelling	Supports multimodal chemistry instruction	Modification	[1]

2 Machine Learning and Data-Driven Learning in Chemistry Education

Within the longer trajectory shown in Figure 1, machine learning marks a shift towards students working directly with chemical data, models and predictions. In many cases, the entry point is not machine learning itself, but a familiar chemical problem. Regression can be introduced through molecular properties, classification through spectra, and predictive modelling through experimental or calculated datasets. Machine learning therefore becomes part of the chemistry activity instead of a separate computational subject.

The reviewed studies show that this approach can be adapted to very different educational settings and technical backgrounds. Some activities require little or no programming, while others introduce students to full scientific-programming workflows. The chemistry

likewise ranges from accessible classification problems to research-oriented molecular and materials modelling. Table 2 gives representative examples of this range.

Table 2: *Representative uses of machine learning and data-driven methods in chemistry education across different educational contexts.*

Educational context	Chemical context	Role of machine learning	Source
Secondary school	Crystal structures and colour chemistry	Introduces classification and prediction through accessible, partly codeless activities	[39, 45]
Introductory university	Molecular and physicochemical properties	Introduces data processing, regression and visualisation through familiar chemical datasets	[43, 46]
Undergraduate	Spectroscopy and chemical data analysis	Uses classification, clustering and pattern recognition to interpret chemical data	[27, 54, 69, 71, 79, 89]
Undergraduate research	Reaction and process modelling	Uses predictive models to study reaction behaviour and limited experimental datasets	[48, 93]
Tool-mediated learning	Chemical data exploration, structure analysis and literature discovery	Allows students to use data-driven tools without implementing the underlying algorithms themselves	[24, 76, 82]
Advanced and research-oriented	Molecular and materials modelling	Connects machine learning with scientific programming, property prediction and contemporary research workflows	[4, 33, 35, 104, 105, 113]

These examples show that machine learning can enter chemistry education at very different technical levels. Students may interact with a guided tool, work in a notebook environment, or construct and evaluate models themselves. At the more advanced end, the distinction between learning machine learning and learning chemistry with machine learning becomes less clear, as data handling, model construction and prediction begin to resemble research practice. This broader shift has been linked to computational thinking, metacognition and changing expectations around data and scientific practice in chemistry [25, 100]. Earlier proposals for integrated AI assistants such as "Molly" anticipated closer interaction between chemical knowledge, computation and intelligent systems [51]. Generative AI has since extended this interaction into more open-ended forms of problem solving [70], which is the focus of the next section.

Despite this technical range, the pedagogical change remains more limited. Most data-driven and machine-learning activities still fit within substitution or augmentation on the SAMR scale. They make chemical analysis, modelling and data exploration more accessible, efficient or authentic, but generally preserve the underlying learning task. Their educational value therefore depends less on the presence of machine learning itself than on whether it helps students reason more effectively with chemical evidence, models and data.

3 Generative AI and Large Language Models in Chemistry Education

Generative AI changes the way students can interact with computational tools. Instead of working mainly with a predefined dataset, model or sequence of exercises, students can ask questions, request explanations, generate text and images, interpret chemical information and use the same system during open-ended problem solving. This flexibility makes large language models educationally attractive, but it also creates a central difficulty. Chemically incorrect responses can be expressed with the same fluency and confidence as correct ones [20, 36, 107]. Their educational value therefore depends not only on what they can produce, but also on whether students and teachers can judge the quality of those outputs.

The reported applications are already broad, ranging from explanation and laboratory writing to course-grounded tutoring, multimodal interpretation, collaborative learning, programming and research support. Table 3 gives representative examples.

The openness of the interaction distinguishes these systems from many earlier forms of educational AI. A student can move from asking for an explanation to questioning a result, revising code or discussing a possible mechanism within the same system. Retrieval-augmented generation can constrain that openness by grounding responses in selected course materials. Professor Leodar supports computational data-science learning in materials chemistry, while ChEdu combines retrieval-augmented generation with guided Socratic questioning [92, 102]. At the same time, greater flexibility does not guarantee better educational design. Chemistry-focused GPTs have shown unnecessary features, misleading descriptions of systems as “experts”, and concerns around copyrighted material[34, 80]. Performance also varies considerably across models and tasks [15, 61, 63, 83, 112]. Evaluation of the system therefore becomes part of using it educationally.

3.1 Capabilities and Accuracy of LLMs in Chemistry

Fluent language does not necessarily reflect sound chemical understanding. Chatbot explanations can reproduce misconceptions and incomplete reasoning patterns similar to those found in students [86, 103, 106]. Yik and Dood found the same problem in 255 ChatGPT explanations of organic reaction mechanisms, many of which appeared sophisticated while still containing chemical inaccuracies [107]. Errors have also been reported in calculations, chemical concepts and references [94]. The limitations become more ap-

Table 3: *Representative uses of generative AI and large language models in chemistry education.*

Educational use	What the system supports	Source
Explanation and writing	Chemical explanations, laboratory-report text and responses to chemistry problems	[20, 36, 107]
Multimodal chemistry	Interpretation of spectra, molecular structures and other forms of visual chemical information	[3, 78]
Course-grounded tutoring	Guided dialogue and responses grounded in instructor or course material	[19, 92, 101, 102]
Collaborative and critical learning	Chatbot-supported discussion, simulated peer review, stories and analogies	[28, 64, 96]
Research and computational support	Coding, modelling, hypothesis testing, literature work and reaction-mechanism reasoning	[16, 26, 31, 60]
Immersive and automated activities	Virtual laboratories, interactive guidance and increasingly automated experimental work	[11, 68, 88]

parent when tasks require application and reasoning. Across 200 high-school chemistry questions, ChatGPT and Microsoft Bing Chat performed less well on questions requiring higher levels of reasoning [15], while performance also depended on how questions were constructed [61]. In acid–base equilibrium problems, students and ChatGPT showed different error patterns: student errors often involved problem conceptualisation and misuse of heuristics, whereas ChatGPT generally conceptualised the problems appropriately but sometimes made mathematical or strategy-related errors [13].

Specialised chemical tasks show a similar pattern. Sharlin and colleagues evaluated seven models on 115 NMR structure-elucidation problems and found that newer reasoning models and greater reasoning effort performed better, whereas structured prompting did not produce consistent improvements. The strongest reasoning model exceeded undergraduate student performance on this particular benchmark, although the models still struggled to integrate some spectral features into a single chemical interpretation [78]. Current systems can therefore provide useful support while remaining vulnerable to plausible but incorrect reasoning [19, 34]. The cumulative insight from these studies is not that generative AI is simply unreliable, but that its usefulness depends strongly on whether users have enough chemical judgement to question, verify and, when necessary, reject its responses.

3.2 Prompt Engineering and Instructional Strategies

How students interact with a model also affects the quality of its responses. Pre-service chemistry teachers who initially relied on simple copy-and-paste prompting improved

both their prompts and their satisfaction with the responses after structured instruction using the "Five S" framework [87]. The requirements vary with the task. Visual generation and interpretation were particularly sensitive to the amount of structure and detail supplied in prompts [65], while explicit prompting strategies in organic chemistry improved the plausibility of model-generated reaction products compared with unguided queries [91]. Prompting should therefore be taught as a task-dependent skill that includes supplying relevant context, checking the response and recognising when the model should not be trusted.

4 AI-Assisted Assessment and Feedback

AI-assisted assessment in chemistry has developed from early attempts to model how students solve problems to systems that can score written explanations, analyse reasoning patterns and support reflection. Across these approaches, the basic promise is that AI can identify patterns in student work more quickly and consistently than manual analysis alone. The critical educational question is what happens after that analysis. Assessment becomes formative only when the information produced changes subsequent teaching, feedback or learning [7]. Figure 3 represents this distinction by placing automated analysis inside a wider cycle of instructor judgement and student response.

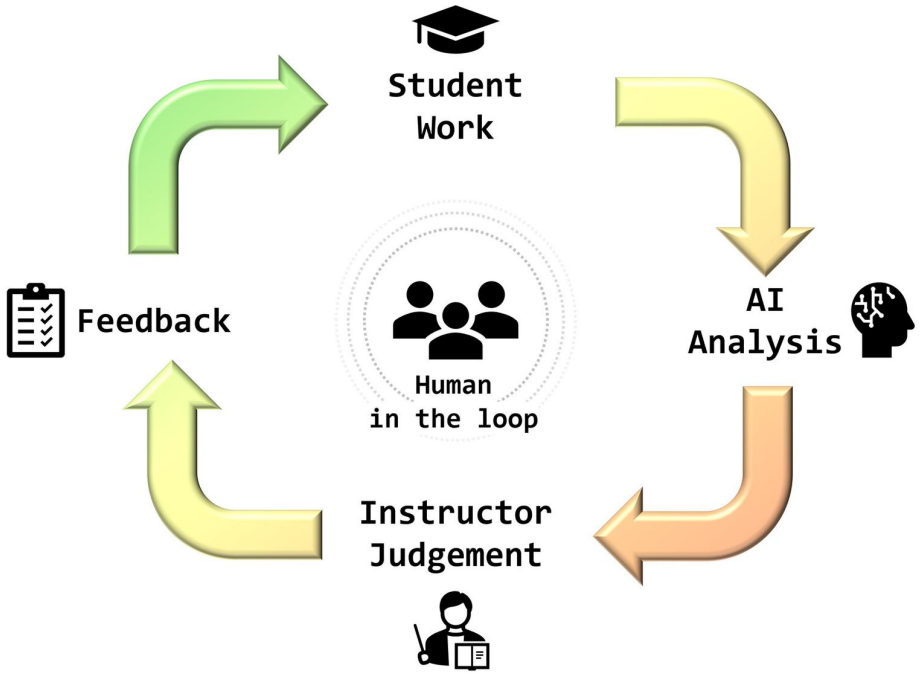


Figure 3: *AI-assisted assessment as a feedback cycle. Automated analysis can identify patterns in student responses, while instructor interpretation and subsequent feedback determine how that information is used for learning.*

The applications reviewed here span several forms of assessment, from cognitive tutoring

and automated grading to analysis of open-ended reasoning and AI-supported reflection. Table 4 gives representative examples.

Table 4: *Representative uses of AI in chemistry assessment and feedback.*

Assessment use	What AI contributes	Source
Cognitive tutoring	Models problem-solving processes and generates worked guidance from chemical principles	[6, 40, 72, 99]
Automated grading	Scores student responses and handles unexpected errors with performance approaching human graders	[41]
Open-ended response analysis	Classifies explanations, hypotheses, argumentation and other forms of written chemical reasoning	[18, 53, 56, 67, 81]
AI-use detection	Uses response-pattern modelling to identify possible generative-AI use in assessments	[83]
Reflection and confidence	Supports examination reflection, confidence calibration and responsible use of generative AI	[2, 98]
Learning analytics and governance	Captures student activity for analysis and frames AI use through human-centred educational principles	[47, 73]

The early work already illustrates an important distinction between analysing student thinking and simply checking whether an answer is correct. Birk used expert systems to examine approaches to chemical reactivity problems [6], while later cognitive tutors generated worked solutions for balancing chemical equations from chemical principles rather than relying on a fixed answer bank [40]. This approach drew on a broader cognitive-tutoring tradition developed extensively in mathematics education [72], and a controlled evaluation in first-semester general chemistry reported significant improvements in performance [99]. The same research tradition later produced an automated grading system capable of handling unexpected student errors with agreement comparable to human graders [41]. In TPACK terms, such systems combine technological, pedagogical and chemistry-specific knowledge [62], although grading itself still largely represents augmentation on the SAMR scale [29].

Later work extended automated assessment to forms of reasoning that cannot be judged as simply correct or incorrect. One study trained classifiers on 1,730 human-coded responses to evaluate explanations of London dispersion forces [67]. Related approaches have been used to classify hypotheses in organic chemistry, score performance-based items, analyse written justifications and evaluate chemical argumentation [18, 53, 56, 81]. These systems make it possible to analyse forms of student reasoning that were previously difficult to score automatically, but they still depend heavily on human-coded examples and decisions

about what counts as a good explanation or argument. Agreement with a human grader also does not by itself demonstrate better learning. Faster and more consistent scoring can support formative assessment, but its educational value depends on what is done with the resulting information.

More recent work has therefore moved beyond scoring alone. AI has been used to support structured reflection before and after examinations [98], identify possible generative-AI use through response-pattern modelling [83], and explore whether students' confidence in AI-assisted answers matches their actual correctness. In one study, 101 undergraduate pharmacy students worked through medicinal-chemistry cases with optional generative-AI assistance, allowing confidence calibration to be examined alongside engagement [2]. At the same time, platforms such as I3Learn can capture detailed behavioural data for analysis [73], while human-centred approaches emphasise teacher judgement, transparency and the risk of "epistemic drift" when opaque systems separate answers from the reasoning that supports them [47]. AI can therefore extend what instructors are able to observe and analyse, but the important issue is how that analysis is turned into feedback that improves chemical reasoning.

5 Personalised and Adaptive Learning Systems

Personalised learning has been an ambition of AI in chemistry education from the beginning. Early systems could respond to an individual question, while later systems began to use previous performance, course context or dialogue history to shape what happens next. The key difference is whether the system only responds to the learner in the moment or adapts to their developing understanding over time. Table 5 summarises representative approaches.

GEORGE provides a useful historical baseline. It acted as a "problem-solving partner" that allowed students to enter their own chemistry problems and receive explanations in natural language [14]. The response was personalised to the immediate question, but the system did not keep track of what a student had previously attempted, misunderstood or asked. Later systems moved further by using information about prior performance. ALEKS, for example, adjusts subsequent content according to what a learner appears to know. This is a more substantial form of adaptation, but not necessarily a better educational outcome. A controlled pretest–posttest study found that students' self-regulated learning skills declined over a semester of ALEKS use [32], while a more recent equity-focused study found that machine-learning-based adaptive support maintained and partly exacerbated educational debts in undergraduate chemistry [57]. Adaptivity therefore depends as much on design as on technical capability.

Conversational systems extend personalisation in another direction. ChEdu combines retrieval-augmented generation with guided Socratic questioning, while Professor Leodar grounds its responses in course materials for computational data-science learning in materials chemistry [92, 102]. Other systems deliberately restrict interaction to instructor-approved content or use conversational strategies to support engagement [38, 49, 50]. Their effectiveness still depends partly on how students use and judge them. Secondary-school students have shown attitudes ranging from acceptance to scepticism [66], and

Table 5: *Representative approaches to personalised and adaptive learning in chemistry education.*

Approach	Form of personalisation or adaptation	Source
Early problem-solving partner	Provides explanations in response to student-entered chemistry problems, but does not build a continuing model of the learner	[14]
Performance-based adaptive learning	Adjusts subsequent content according to a learner's previous performance	[32, 57]
Course-grounded conversational tutors	Uses course materials and guided dialogue to provide context-sensitive support	[92, 102]
Curated and engagement-oriented chatbots	Restricts responses to approved content or uses conversational strategies to support engagement	[38, 49, 50]
Open-ended scaffolding	Supports students while they retrieve information, critique models and work through less structured problems	[31, 70]
Visual and immersive support	Adapts the form of explanation through narrative, visual representation or interactive virtual environments	[1, 88]

students differ in their ability to assess whether chatbot responses are trustworthy [109]. ChEdu and Professor Leodar are among the few examples in this review that plausibly reach the Modification level of SAMR because the learning activity shifts from retrieving a fixed answer to participating in a guided dialogue. They still fall short of Redefinition, since the learning goals and boundaries remain largely defined by existing course materials.

Generative AI also broadens what personalisation can involve. Students engaged in computational modelling have used it to retrieve and clarify information and to test and critique their own models [31], while similar support has been reported in collaborative industrial-design work in unit operations [70]. Personalisation can also concern how material is represented. An AI-generated graphic story grounded in Universal Design for Learning was used to teach protein structure through narrative and visual representation [1], while AI-powered avatars in virtual-reality laboratory environments provide another form of personalised guidance [88]. These approaches extend adaptation beyond selecting the next task to changing how students encounter and work with chemical information.

Across these systems, responsiveness is much more common than genuine adaptation. Many tools can answer an individual question or alter the form of a response, but far fewer have been shown to build a reliable picture of a learner's developing understanding and use it to improve learning over time. The long-term evidence remains limited, and the ALEKS findings show that adaptation can also have undesirable effects. Personalised learning should therefore be judged by whether it supports stronger reasoning, greater independence and more durable chemical understanding, not simply by how tailored the interface appears.

6 AI in Laboratory Instruction and Simulation

Artificial intelligence is entering chemistry laboratory education at several points in the experimental process. It can support preparation and reporting, guide students through virtual environments, assist with experimental automation, and analyse or visualise laboratory data. These applications differ substantially in educational value. Some mainly make an existing task easier, while others allow students to engage with modelling, automation and experimental decision-making in ways that would otherwise be difficult to incorporate into a teaching laboratory. Table 6 summarises representative examples.

6.1 Generative AI in Lab Preparation and Reporting

The most immediate use of generative AI in laboratory courses is before and after the experiment. Large language models can help students draft laboratory discussions or suggest possible modifications to experimental procedures. The same flexibility creates an obvious limitation. ChatGPT can generate plausible laboratory-report text that contains subtle errors [36], and suggestions for modifying experiments can be vague or chemically unsuitable even when they appear reasonable [74]. In this setting, the educational value lies less in generating text or ideas than in requiring students to judge them against

Table 6: *Representative uses of AI and data-driven methods in chemistry laboratory education.*

Laboratory use	What AI or computation contributes	Source
Preparation and reporting	Supports laboratory writing, planning and development of experimental ideas	[36, 74]
Modelling and automation	Brings predictive modelling, experimental optimisation and automated liquid handling into laboratory activities	[68, 93]
Virtual and immersive laboratories	Provides interactive guidance, procedural support and safety instruction in simulated laboratory environments	[11, 82, 88]
Simulation and digital data	Connects molecular simulation and machine-readable experimental data with computational analysis	[35, 75, 104]
Experimental data analysis	Uses classification, regression and generative tools to analyse, interpret and visualise experimental measurements	[43, 84, 89, 90]

chemical knowledge and experimental constraints.

More substantial changes occur when students work directly with computational models or automation. In an inquiry-based dry laboratory, students constructed molecular descriptors and used regression to predict Diels–Alder reaction barriers [93]. A first-year activity went further by asking students to assemble a low-cost automated liquid-handling workstation and use it to optimise synthesis conditions for a 3d–4f cluster [68]. In such cases, AI and automation become part of experimental reasoning itself rather than an external tool for producing a report.

6.2 Virtual Labs and Simulations

Virtual environments extend this idea by allowing students to interact with experiments and analytical tasks without relying entirely on a physical laboratory. AI-supported systems range from ACD/Structure Elucidator, introduced in Section 2, to virtual-reality laboratories in which conversational agents guide users through procedures and safety decisions [11, 82, 88]. These environments can be especially useful when repeated practice, hazardous procedures or access to equipment would otherwise limit experimentation.

Simulation also provides access to chemical phenomena that cannot easily be explored directly in an instructional laboratory. Machine-learning potentials have been used to construct potential-energy surfaces for hydrogen dissociation on Cu(111) and to examine acidity in zeolite frameworks [104]. Other activities focus on the data infrastructure behind modern experimentation. FAIR-oriented laboratory exercises, for example, expose students to machine-readable experimental metadata that can later support computational

and AI-based analysis [75]. Attempts to integrate similar approaches into polymer chemistry also show that the main difficulty is often curricular and practical rather than purely technical [35].

The SAMR distinction is useful here. Structure-elucidation software and many simulation exercises mainly augment analytical activities that already existed. Virtual-reality guidance and chatbot-mediated 3D laboratories move further because the student can interact with a simulated environment and receive context-sensitive support during the activity. These approaches therefore have greater potential for Modification, although the presence of an immersive interface does not by itself demonstrate deeper learning.

6.3 Data Analysis and Visualisation in Labs

AI also changes what students can do with the data produced in laboratory and computational activities. Classification and regression can be used to interpret spectra or model chemical properties, while generative systems can assist with scientific visualisation [43, 84, 89]. The educationally stronger examples do more than replace conventional data processing. A machine-learning version of the particle-in-a-box experiment using cyanine-dye absorption spectra, for example, asks students to generate and test hypotheses about the underlying physical model instead of simply reproducing a predetermined fit [90].

Across these applications, the important distinction is whether AI removes students from the experimental reasoning process or gives them new ways to participate in it. Automated reporting can save time without changing what students understand. Modelling, simulation and automation become more educationally significant when students must decide what data to collect, how a model should be constructed, whether its output is chemically plausible and what should happen next. The value of AI in laboratory education therefore lies less in automating experimental work than in expanding the forms of experimental reasoning that students can practise.

7 Reliability, Integrity, and Equity

As AI becomes more capable, the educational challenge increasingly shifts from access to judgement. Students and teachers need to decide when an AI-generated response is trustworthy, when AI use changes what an assessment actually measures, and whether the benefits of these systems are distributed equitably. This places new demands on AI literacy. Technological knowledge alone is not enough. Teachers also need prompting skills and an understanding of bias, discrimination and hallucination [23]. Evidence from chemistry educators suggests that these competencies are still developing. Teachers have reported weaknesses in their own AI literacy and in their ability to design lessons that use AI effectively [108]. Interviews with three chemistry faculty found considerable uncertainty about the role of AI in student learning [58], while surveys of 100 trainee chemistry teachers and 285 practising chemistry teachers found moderate or limited AI awareness and continuing uncertainty about classroom application [10, 30].

Some curricular responses already address these needs. AI concepts have been incorporated directly into secondary chemistry teaching [59], while a graduate biochemistry course used several generative-AI platforms to develop understanding of transformer-based systems alongside chemistry content [77]. The aim is not simply to teach students how to use AI, but to connect its use with chemical reasoning, critical evaluation and responsible scientific practice [9, 19].

The difficulty is that the technology continues to change faster than the evidence base. The capabilities and uses of generative AI have shifted substantially within only a few years [110, 112], while its educational benefits remain less certain [34]. Other issues receive even less attention. Student prompts, mistakes and learning behaviour may be processed by external providers, raising questions about privacy, while large language models also carry environmental costs associated with substantial energy use and carbon emissions [52]. These concerns provide the broader context for the reliability, integrity and equity issues considered below and mapped in Figure 4.

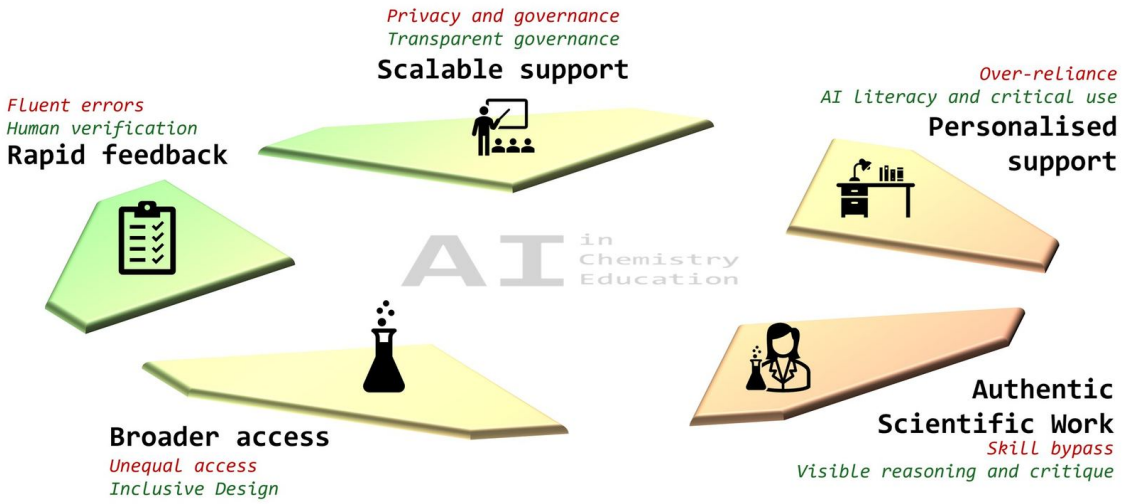


Figure 4: *Opportunities, risks and pedagogical responses associated with AI in chemistry education. Productive use depends on evidence, human oversight, critical AI literacy and equitable access.*

7.1 Reliability as an Educational Risk

As discussed in Section 3.1, generative models can produce chemically plausible responses that nevertheless contain conceptual or factual errors. The educational risk is greatest when students cannot distinguish a convincing explanation from a chemically sound one. This is especially important for novices. Cognitive load theory helps explain why [85]. Experienced chemists can often recognise an implausible mechanism or inconsistent explanation because they already possess organised chemical knowledge, while students may lack exactly the knowledge required to detect the same error. Adversarial testing reinforces this concern by showing that models can be led into confidently stated chemical errors even on material they otherwise handle successfully [95]. Reliability

therefore cannot be addressed only by improving model performance. Students also need explicit practice in challenging, verifying and rejecting AI-generated chemical reasoning. Verification should therefore be treated as part of learning chemistry with AI, not as an additional digital skill.

7.2 Academic Integrity and Assessment Design

Generative AI changes what assessment can be assumed to measure. As discussed in Section 3, AI-generated responses can sometimes be distinguished statistically from student work, but detection does not resolve the deeper problem. If a system can generate a plausible explanation, report or solution, an assessment based only on the final product may no longer reveal how the student arrived there.

A more durable response is to design assessments that make student reasoning visible through interpretation, revision and justification. Practical guidance for chemistry teachers increasingly emphasises transparency, discussion of AI limitations and responsible use rather than prohibition alone [8]. Responsibility also extends to researchers and educators, who should disclose the role of generative AI and remain alert to its limitations and biases [17]. Academic integrity is therefore not only about whether AI was used, but whether the assessment still provides evidence of the student's own chemical reasoning.

7.3 Equity and Access

Equitable access involves more than whether students can open the same chatbot. Disability and neurodivergence, language, geography, digital infrastructure and the ability to evaluate AI output all shape who benefits from these systems. Disability and neurodivergence are especially weakly represented in chemistry-education research. Work in STEM education more broadly suggests that AI could support personalised learning when combined with Universal Design for Learning principles, while also raising concerns about algorithmic bias and student data privacy [44]. An important earlier chemistry-specific exception developed an AI-based chemistry tutor for blind and visually impaired students and designed it for screen-reader accessibility [42]. However, contemporary implementation and outcome evidence for disabled and neurodivergent learners remains sparse. A professional-development model for AI-mediated inclusive chemistry teaching begins to address this problem at the level of teacher preparation, but does not yet report outcomes for the students it is intended to support [22]. Evidence from actual chemistry classrooms therefore remains a clear research need.

Language and geography create a different form of inequality. Locally deployed language models have been explored in Chinese high schools as a response to shortages in teaching support and unequal access to educational resources [101]. A multilingual machine-learning system for postsecondary chemistry takes another approach by allowing students who are still learning the language of instruction to engage with formative assessment in a language they command [55]. These examples suggest that AI can reduce some barriers when localisation and linguistic diversity are treated as design requirements rather than afterthoughts.

Technical literacy creates another divide. Access to an AI system does not mean that a learner can judge whether it is trustworthy. Chemistry-focused GPTs have been found to include unnecessary features and misleading descriptions using terms such as "expert" and "teacher" [80]. Students with less disciplinary or technical knowledge may therefore be more easily misled by the appearance of authority. Equity also concerns who participates in developing and evaluating educational AI. Generative AI has been proposed as one way of supporting researchers with disabilities and broadening participation in chemistry-education research [17]. The issue is therefore not only who can access AI, but who can evaluate it, shape it and benefit from its use.

Across reliability, integrity and equity, the same principle recurs. Greater technical capability does not remove the need for human judgement. In many cases, it increases it. Responsible use of AI in chemistry education therefore depends on developing the chemical knowledge, AI literacy and educational structures needed to question these systems, determine when their use is appropriate and ensure that their benefits do not remain limited to already advantaged learners.

8 Conclusions and Outlook

Four decades of AI in chemistry education show that technical capability has developed much faster than educational practice. From GEORGE to current generative AI, the range of tasks that can be supported or automated has expanded considerably [14, 111]. Yet most applications reviewed here still fall within substitution or augmentation, while convincing examples of Redefinition remain rare [29, 62]. This does not reduce the importance of AI in chemistry education. It means that its value depends on how it is used. Students are increasingly expected to combine chemical knowledge with computation, data handling, automation and reproducible workflows [9], while teachers need to decide where AI genuinely supports learning and where it simply makes an existing task faster.

The evidence is still limited in several important ways. Most studies concern individual courses or short interventions, and there is little evidence that reported gains persist after the tool is removed or transfer to unfamiliar chemical problems [63, 112]. Open and locally deployable systems, specialised chemistry GPTs and immersive environments also need to be compared more directly in terms of learning outcomes, chemical accuracy, usability, cost, privacy and accessibility [1, 19, 80, 88]. These priorities extend from better evaluation of current tools to longer-term questions about how AI should be integrated into chemistry teaching and research, as outlined in Figure 5.

The stakes will increase as AI moves from explaining chemistry to participating more directly in chemical work. Language models are already being used to support mechanism analysis, troubleshooting and experimental reasoning [16]. Recommending an experiment or synthesis route, however, carries different consequences from explaining one after the fact. As these systems become more involved in scientific practice, chemical judgement, human oversight and clear standards for their use will matter more, not less [34, 47]. The next phase of AI in chemistry education should therefore be shaped by what students actually learn from these systems and by how well they are prepared to use them critically in real chemical practice.

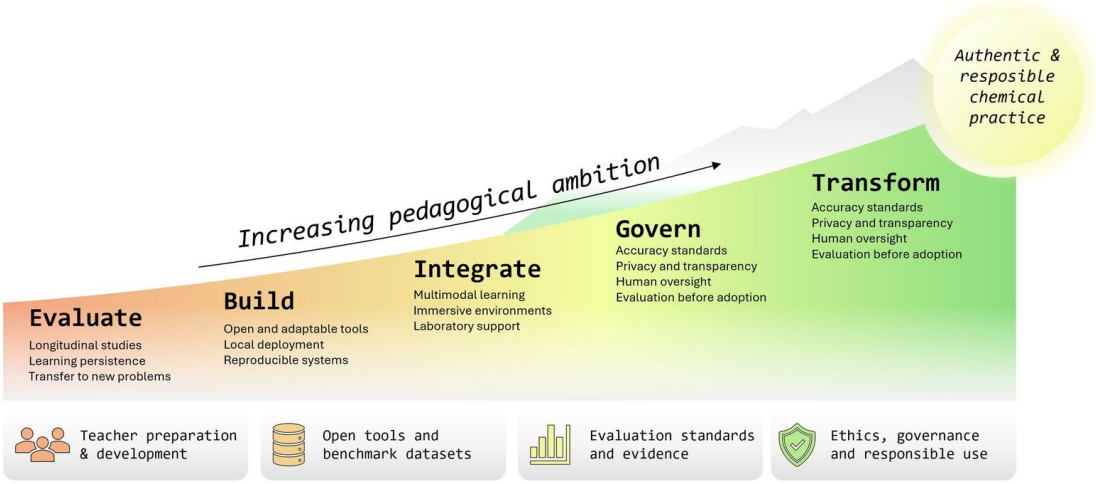


Figure 5: *Roadmap for AI in chemistry education across near-, medium- and long-term priorities, supported by teacher preparation, benchmark datasets, evaluation standards and governance.*

Author contributions

Aleksandar Kondinski is the sole author and was responsible for conceptualisation, literature review, analysis, manuscript drafting, figure preparation, and final approval of the manuscript.

Conflicts of interest

The author declares no competing interests.

Data availability

No new data are associated with this article.

Acknowledgements

No external funding supported the underlying research. The author thanks the TU Graz Open Access Publishing Fund for support. OpenAI ChatGPT was used to assist with language editing, grammar, and clarity; the author reviewed and edited the manuscript and takes full responsibility for its content.

References

[1] S. Aghasafari, L. Cai, and M. Malloy. Visualizing protein structure: Integrating ai, graphic storytelling, and media arts in steam-based chemistry education. *Journal of Chemical Education*, 103(3):1214–1224, 2026. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00586.

[2] A. H. Al-Nadaf. Generative artificial intelligence in medicinal chemistry education: Student engagement, confidence calibration, and safety-related uncertainty in case-based learning. *Journal of Chemical Education*, 103(9):4949–4958, 2026. doi:10.1021/acs.jchemed.6c00013.

[3] E. A. Alasadi and C. R. Baiz. Multimodal generative artificial intelligence tackles visual problems in chemistry. *Journal of Chemical Education*, 101(7):2716–2729, 2024. ISSN 0021-9584. doi:10.1021/acs.jchemed.4c00138.

[4] M. Amrihesari and B. Brettmann. Introducing data-driven materials informatics into undergraduate courses through a polymer science workshop. *Journal of Chemical Education*, 102(9):3972–3981, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00562.

[5] S. Berber, M. Brueckner, N. Maurer, and J. Huwer. Artificial intelligence in chemistry research-implications for teaching and learning. *Journal of Chemical Education*, 102(4):1445–1456, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.4c01033.

[6] J. P. Birk. The computer as student: an application of artificial intelligence. *Journal of Chemical Education*, 69(4):294, 1992. ISSN 0021-9584. doi:10.1021/ed069p294.

[7] P. Black and D. Wiliam. Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1):7–74, 1998. doi:10.1080/0969595980050102.

[8] R. Blonder and Y. Feldman-Maggor. Ai for chemistry teaching: responsible ai and ethical considerations. *Chemistry Teacher International*, 6(4):385–395, 2024. doi:10.1515/cti-2024-0014.

[9] S. Chakraborty. From beakers to bytes: educating the hybrid chemists - a chemistry education perspective on integrating wet and dry labs: a call for hybrid skills in the next generation of chemist. *ChemRxiv*, pages 1–20, 2025. ISSN 2573-2293. doi:10.26434/chemrxiv-2025-4b720.

[10] T. S. Chandran and Y. Mhd Bakri. Awareness of generative artificial intelligence in chemistry education among trainee teachers. *Journal of Education and Learning (EduLearn)*, 20(1):424–430, 2026. doi:10.11591/edulearn.v20i1.23005.

[11] C.-M. Chen, B.-J. Chen, and C.-L. Huang. An enhanced ai chatbot to facilitate learning performance in a virtual 3d chemistry laboratory. *Journal of Science Education and Technology*, 2026. doi:10.1007/s10956-026-10319-3.

[12] W.-K. Chiu. Pedagogy of emerging technologies in chemical education during the era of digitalization and artificial intelligence: A systematic review. *Education Sciences*, 11(11):709, 2021. doi:10.3390/educsci11110709.

[13] T. M. Clark, E. Anderson, N. M. Dickson-Karn, C. Soltanirad, and N. Tafini. Comparing the performance of college chemistry students with chatgpt for calculations involving acids and bases. *Journal of Chemical Education*, 100(10):3934–3944, 2023. doi:10.1021/acs.jchemed.3c00500.

[14] R. Cornelius, D. Cabrol, and C. Cachet. Applying the techniques of artificial intelligence to chemistry education. *ACS Symposium Series*, 306:125, 1986. ISSN 0097-6156. doi:10.1021/bk-1986-0306.ch011.

[15] X.-Q. Dao, N.-B. Le, T.-D. Vo, B.-B. Ngo, and X.-D. Phan. Llms' capabilities at the high school level in chemistry: cases of chatgpt and microsoft bing chat. *ChemRxiv*, pages 1–7, 2023. ISSN 2573-2293. doi:10.26434/chemrxiv-2023-kxxpd.

[16] A. Das, S. Behera, R. Haldar, P. Dawar, S. Maity, and S. De. Generative ai as chemists virtual assistant: a case study with the chemistry of smi2. *Discover Applied Sciences*, 8(2):166, 2026. ISSN 3004-9261. doi:10.1007/s42452-025-07913-2.

[17] J. M. Deng, Z. Lalani, L. A. McDermaid, and A. R. Szozda. Using generative artificial intelligence in chemistry education research: prioritizing ethical use and accessibility. *ChemRxiv*, pages 1–22, 2023. ISSN 2573-2293. doi:10.26434/chemrxiv-2023-24zfl-v2.

[18] A. J. Dood, F. M. Watts, M. C. Connor, and G. V. Shultz. Automated text analysis of organic chemistry students' written hypotheses. *Journal of Chemical Education*, 101(3):807–818, 2024. ISSN 0021-9584. doi:10.1021/acs.jchemed.3c00757.

[19] Y. Du, C. Duan, A. M. Bran, A. Sotnikova, Y. Qu, H. Kulik, A. Bosselut, J. Xu, and P. Schwaller. Large language models are catalyzing chemistry education. *ChemRxiv*, pages 1–20, 2024. ISSN 2573-2293. doi:10.26434/chemrxiv-2024-h722v.

[20] M. E. Emenike and B. U. Emenike. Was this title generated by chatgpt considerations for artificial intelligence text-generation software programs for chemists and chemistry educators. *Journal of Chemical Education*, 100(4):1413–1418, 2023. ISSN 0021-9584. doi:10.1021/acs.jchemed.3c00063.

[21] A. K. Erümit and R. Özdemir Sarıalıoğlu. Artificial intelligence in science and chemistry education: A systematic review. *Discover Education*, 4(1), 2025. doi:10.1007/s44217-025-00622-3.

[22] P. A. Estrada Palencia, K. P. Agudelo Arteaga, and E. P. Flórez Nisperuza. Artificial intelligence and inclusive education: A professional development model for chemistry teachers in secondary education. *Educational Process International Journal*, 17(1), 2025. doi:10.22521/edupij.2025.17.380.

[23] Y. Feldman-Maggor, R. Blonder, and G. Alexandron. Perspectives of generative ai in chemistry education within the tpack framework. *Journal of Science Education and Technology*, 34(1):1–12, 2024. doi:10.1007/s10956-024-10147-3.

[24] J.-P. Francoia and L. Vial. Chembrows: An open-source application software to keep up to date with the current literature. *Journal of Chemical Education*, 93(6): 1137–1139, 2016. ISSN 0021-9584. doi:10.1021/acs.jchemed.6b00024.

[25] F. M. Fung, M. Lederbauer, Y. S. L. Choo, T. Gehring, K. M. Jablonka, K. Jorner, P. Schwaller, M. B. Sullivan, A. Volkamer, M. S. Sigman, K. Liao, and C. Windle. Chemical education in digital chemistry. *Chem*, 10(12):3519–3525, 2024. ISSN 2451-9294. doi:10.1016/j.chempr.2024.10.010.

[26] C. K. Gordon, L. D. Browne, T. G. Sweet, C. Zemke-Smith, F. J. Blaauw-Smith, S. V. Chan, M. Waterland, and N. J. L. K. Davis. Enhancing student confidence in data processing through chatgpt-assisted coding in chemistry education. *Discover Education*, 5(1), 2026. doi:10.1007/s44217-026-01479-w.

[27] A. Grant St James, L. Hand, T. Mills, L. Song, A. S. J. Brunt, P. E. Bergstrom Mann, A. F. Worrall, M. I. Stewart, and C. Vallance. Exploring machine learning in chemistry through the classification of spectra: An undergraduate project. *Journal of Chemical Education*, 100(3):1343–1350, 2023. doi:10.1021/acs.jchemed.2c00682.

[28] K. Hallal, R. Hamdan, and S. Tlais. Beyond traditional texts: Exploring ai as a tool for generating stories and analogies in organic chemistry education. *Journal of Chemical Education*, 103(2):846–857, 2026. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00693.

[29] E. R. Hamilton, J. M. Rosenberg, and M. Akcaoglu. The substitution augmentation modification redefinition (samr) model: A critical review and suggestions for its use. *TechTrends*, 60(5):433–441, 2016. doi:10.1007/s11528-016-0091-y.

[30] R. Hammoud. Exploring chemistry teachers' knowledge and perceptions of ai: Opportunities and challenges. *International Journal of Research in Education and Science*, 11(1):49–60, 2025. doi:10.46328/ijres.3584.

[31] A. Haraldsrud and T. O. B. Odden. Bridging artificial intelligence and real intelligence: Self-scaffolding computational modeling with generative ai in chemistry. *Journal of Chemical Education*, 102(10):4255–4266, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00657.

[32] H. Harati, L. Sujo-Montes, C.-H. Tu, S. Armfield, and C.-J. Yen. Assessment and learning in knowledge spaces (aleks) adaptive system impact on students' perception and self-regulated learning skills. *Education Sciences*, 11(10):603, 2021. doi:10.3390/educsci11100603.

[33] J. Heras-Domingo and D. Garay-Ruiz. Pythonic chemistry: The beginner's guide to digital chemistry. *Journal of Chemical Education*, 101(11):4883–4891, 2024. ISSN 0021-9584. doi:10.1021/acs.jchemed.4c00840.

[34] G. M. Hocky and A. D. White. Looking back and to the future after four-plus years of language in chemistry. *Digital discovery*, 5(4):1440–1446, 2026. ISSN 2635-098X. doi:10.1039/d5dd00550g.

[35] H. Hu, C. Liu, L. Guo, H. Yuan, and L. Huang. Application of artificial intelligence in polymer chemistry teaching: innovation, practice and implicit challenges. *Daxue Huaxue*, 40(9):178–188, 2025. ISSN 1000-8438. doi:10.12461/pku.dxhx202503010.

[36] T. Humphry and A. L. Fuller. Potential chatgpt use in undergraduate chemistry laboratories. *Journal of Chemical Education*, 100(4):1434–1436, 2023. ISSN 0021-9584. doi:10.1021/acs.jchemed.3c00006.

[37] A. Iyamuremye, F. N. Niyonzima, J. Mukiza, I. Twagilimana, P. Nyirahabimana, T. Nsengimana, J. D. Habiyaremye, O. Habimana, and E. Nsabayezu. Utilization of artificial intelligence and machine learning in chemistry education: A critical review. *Discover Education*, 3(1), 2024. doi:10.1007/s44217-024-00197-5.

[38] J. Jasin, H. T. Ng, I. Atmosukarto, P. Iyer, F. Osman, P. Y. K. Wong, C. Y. Pua, and W. S. Cheow. The implementation of chatbot-mediated immediacy for synchronous communication in an online chemistry course. *Education and Information Technologies*, 28(8):10665–10690, 2023. doi:10.1007/s10639-023-11602-1.

[39] S. Jiang, J. McClure, H. Mao, J. Chen, Y. Liu, and Y. Zhang. Integrating machine learning and color chemistry: Developing a high-school curriculum toward real-world problem-solving. *Journal of Chemical Education*, 101(2):675–681, 2024. doi:10.1021/acs.jchemed.3c00589.

[40] B. G. Johnson and D. A. Holder. A cognitive modeling tutor supporting student inquiry for balancing chemical equations. *Chemical Educator [online computer file]*, 7(5):297–302, 2002. ISSN 1430-4171. doi:10.1007/s00897020592a.

[41] B. G. Johnson and D. A. Holder. An artificial intelligence-based program for automated grading of student work in balancing chemical equations. *Chemical Educator*, 8(4):271–279, 2003. ISSN 1430-4171. doi:10.1333/s00897030703a.

[42] B. G. Johnson, J. S. Dittel, and D. A. Holder. Accessible artificial intelligence-based chemistry tutoring for blind and visually impaired students. *Chemical Educator*, 15:171–177, 2010. ISSN 1430-4171. doi:10.1333/s00897102271a.

[43] L. Joss and E. A. Muller. Machine learning for fluid property correlations: Classroom examples with matlab. *Journal of Chemical Education*, 96(4):697–703, 2019. ISSN 0021-9584. doi:10.1021/acs.jchemed.8b00692.

[44] S. Kohnke and T. Zaugg. Artificial intelligence: An untapped opportunity for equity and access in stem education. *Education Sciences*, 15(1):68, 2025. doi:10.3390/educsci15010068.

[45] T. Kraska. Classification of nacl and cscl crystal structures: Establishing an integrated didactic approach for machine learning and chemistry in upper secondary education. *Journal of Chemical Education*, 102(11):4968–4975, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00209.

[46] D. Lafuente, B. Cohen, G. Fiorini, A. A. Garcia, M. Bringas, E. Morzan, and D. Onna. A gentle introduction to machine learning for chemists: An undergraduate workshop using python notebooks for visualization, data processing, analysis, and modeling. *Journal of Chemical Education*, 98(9):2892–2898, 2021. ISSN 0021-9584. doi:10.1021/acs.jchemed.1c00142.

[47] Y. Li, L. Tolosa, F. Rivas Echeverria, and R. Marquez. Translating unesco artificial intelligence guidelines to chemical education and its intersection with sustainable development goals. *Journal of Chemical Education*, 103(3):1135–1144, 2026. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00819.

[48] Y. Liu, L. Gao, Z. Yang, H. Zhang, J. Chen, J. Xu, and K. Yin. Using response surface methodology to enhance limited data sets for machine learning: A polyvinyl butyral synthesis-based chemistry undergraduate teaching case. *Journal of Chemical Education*, 102(10):4424–4434, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00505.

[49] A. T. Lolinco and T. A. Holme. Developing a curated chatbot as an exploratory communication tool for chemistry learning. *Journal of Chemical Education*, 100 (10):4092–4098, 2023. doi:10.1021/acs.jchemed.3c00520.

[50] A. T. Lolinco and T. A. Holme. Understanding student help-seeking for contextualizing chemistry through curated chatbot data analysis. *Journal of Chemical Education*, 101(11):4837–4846, 2024. doi:10.1021/acs.jchemed.4c00766.

[51] Y. Lu. The "ok, molly" chemistry. *Accounts of Chemical Research*, 50(3):647–651, 2017. ISSN 0001-4842. doi:10.1021/acs.accounts.7b00009.

[52] A. S. Luccioni, S. Viguier, and A.-L. Ligozat. Estimating the carbon footprint of bloom, a 176b parameter language model. *arXiv preprint*, 2022. doi:10.48550/arxiv.2211.02001.

[53] S. Maestrales, X. Zhai, I. Touitou, Q. Baker, B. Schneider, and J. Krajcik. Using machine learning to score multi-dimensional assessments of chemistry and physics. *Journal of Science Education and Technology*, 30(2):239–254, 2021. doi:10.1007/s10956-020-09895-9.

[54] B. Mahjour, A. McGrath, A. Outlaw, R. Zhao, C. Zhang, and T. Cernak. Interactive python notebook modules for chemoinformatics in medicinal chemistry. *Journal of Chemical Education*, 100(12):4895–4902, 2023. ISSN 0021-9584. doi:10.1021/acs.jchemed.3c00357.

[55] P. P. Martin and N. Graulich. Beyond language barriers: Allowing multiple languages in postsecondary chemistry classes through multilingual machine

learning. *Journal of Science Education and Technology*, 33(3):333–348, 2024. doi:10.1007/s10956-023-10087-4.

[56] P. P. Martin, D. Kranz, P. Wulff, and N. Graulich. Exploring new depths: Applying machine learning for the analysis of student argumentation in chemistry. *Journal of Research in Science Teaching*, 61(8):1757–1792, 2024. doi:10.1002/tea.21903.

[57] P. P. Martin, M. Kubsch, B. J. Yik, B. T. Burlingham, and N. Graulich. Adaptive, but equitable? exploring the impact of machine learning-based adaptive support on educational debts in undergraduate chemistry. *Science Education*, 110(3):928–946, 2026. doi:10.1002/sce.70042.

[58] M. B. Matthews and D. W. Potter. Chemistry faculty perspectives on artificial intelligence and pedagogical change in higher education. *Discover Education*, 5 (1), 2026. doi:10.1007/s44217-026-01250-1.

[59] N. Maurer and J. Huwer. Teaching and learning about artificial intelligence (ai) in secondary education chemistry classes. *Journal of Chemical Education*, 103(9): 4794–4804, 2026. doi:10.1021/acs.jchemed.6c00128.

[60] S. C. McAfee and J.-M. G. Rodriguez. I asked chatgpt to do my research: Welcoming artificial intelligence to the chemistry education research team. *Journal of Chemical Education*, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00958.

[61] G. M. McLeod, B. Hanks, H. Ashby, J. Nolder, B. Sorenson, and K. Hanson. Generative ai vs student performance on multiple-choice, general chemistry exams: Insights into question design from regression analysis. *Journal of Chemical Education*, 103(2):872–878, 2026. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00979.

[62] P. Mishra and M. J. Koehler. Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6):1017–1054, 2006. doi:10.1111/j.1467-9620.2006.00684.x.

[63] C. Morbidoni, A. Mascitti, A. Sarra, L. Tonucci, and F. Coccia. Exploring the potential of using chatgpt in chemistry education. *Journal of Chemical Education*, 103(1):286–297, 2026. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00854.

[64] P. J. P. Naeyaert, L. R. J. Scarratt, T. Murphy, and R. Pullen. Think-pair-chatbot-share: Ai-facilitated peer learning in chemistry. *Journal of Chemical Education*, 102(9):3936–3944, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00438.

[65] W. J. D. J. Nascimento, C. Morais, and G. J. Girotto. Enhancing ai responses in chemistry: Integrating text generation, image creation, and image interpretation through different levels of prompts. *Journal of Chemical Education*, 101(9):3767–3779, 2024. ISSN 0021-9584. doi:10.1021/acs.jchemed.4c00230.

[66] L. Naumann, J. L. Mayr, and S. Markic. Between acceptance and scepticism: an investigation into secondary school students' attitudes toward artificial intelligence in chemistry education. *Chimia*, 79(10):693–697, 2025. ISSN 0009-4293. doi:10.2533/chimia.2025.693.

[67] K. Noyes, R. L. McKay, M. Neumann, K. C. Haudek, and M. M. Cooper. Developing computer resources to automate analysis of students' explanations of london dispersion forces. *Journal of Chemical Education*, 97(11):3923–3936, 2020. ISSN 0021-9584. doi:10.1021/acs.jchemed.0c00445.

[68] M.-Q. Qi, Y.-B. Qi, J.-C. Lv, G.-P. Chen, X.-Y. Diao, Z.-A. Cai, M.-H. Du, J.-P. Li, C. Wang, X.-J. Kong, Y. Ren, L. Long, and L. Zheng. From programming to crystal growth: High-throughput exploration of 3d-4f cluster synthesis conditions-a laboratory experiment for first year undergraduates. *Journal of Chemical Education*, 2026. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c01768.

[69] R. Qiu, Z. Lin, Z. Yang, and L. Gao. Machine learning for infrared spectral classification of polyvinyl butyrals with identical chemical groups: An example for undergraduate chemistry classes. *Journal of Chemical Education*, 101(2):328–336, 2024. doi:10.1021/acs.jchemed.3c00213.

[70] B. Ramos and R. Condotta. Enhancing learning and collaboration in a unit operations course: Using ai as a catalyst to create engaging problem-based learning scenarios. *Journal of Chemical Education*, 101(8):3246–3254, 2024. ISSN 0021-9584. doi:10.1021/acs.jchemed.4c00244.

[71] D. Revignas and V. Amendola. Artificial neural networks applied to colorimetric nanosensors: An undergraduate experience tailorable from gold nanoparticles synthesis to optical spectroscopy and machine learning. *Journal of Chemical Education*, 99(5):2112–2120, 2022. doi:10.1021/acs.jchemed.1c01288.

[72] S. Ritter, J. R. Anderson, K. R. Koedinger, and A. T. Corbett. Cognitive tutor: Applied research in mathematics education. *Psychonomic Bulletin & Review*, 14 (2):249–255, 2007. doi:10.3758/bf03194060.

[73] M. Roski, R. Ewerth, A. Hoppe, and A. Nehring. Exploring data mining in chemistry education: Building a web-based learning platform for learning analytics. *Journal of Chemical Education*, 101(3):930–940, 2024. ISSN 0021-9584. doi:10.1021/acs.jchemed.3c00794.

[74] E. F. Ruff, J. L. Franz, and J. K. West. Using chatgpt for method development and green chemistry education in upper-level laboratory courses. *Journal of Chemical Education*, 101(8):3224–3232, 2024. ISSN 0021-9584. doi:10.1021/acs.jchemed.4c00193.

[75] H. S. Rzepa and S. Kuhn. A data-oriented approach to making new molecules as a student experiment: artificial intelligence-enabling fair publication of nmr data for organic esters. *Magnetic Resonance in Chemistry*, 60(1):93–103, 2022. ISSN 0749-1581. doi:10.1002/mrc.5186.

[76] M. Schmuker, F. Schwarte, A. Brueck, E. Proschak, Y. Tanrikulu, A. Givehchi, K. Scheiffele, and G. Schneider. Sommer: self-organising maps for education and research. *Journal of Molecular Modeling*, 13(1):225–228, 2007. ISSN 0948-5023. doi:10.1007/s00894-006-0140-0.

[77] Y. V. Sevryugina, K. Collins-Thompson, and N. G. Walter. Integrating ai literacy in chemistry graduate education: Harnessing the power of transformer-based models. *AI in Education*, 2(2):14, 2026. doi:10.3390/aieduc2020014.

[78] S. Sharlin, F. Agbere, K. Ishimwe, Z. Osifova, O. Socha, M. Dracinsky, and T. R. Josephson. Nmr-challenge for llms: evaluating chemical reasoning in humans and ai. *ChemRxiv*, pages 1–16, 2025. ISSN 2573-2293. doi:10.26434/chemrxiv-2025-x8h36-v2.

[79] D. Shiryaev, Y. Sun, E. I. Jaffal, and A. O. Oliynyk. Exploring feature engineering for crystal structure classification: interactive applications of pca and pls-da clustering. *ChemRxiv*, pages 1–15, 2025. ISSN 2573-2293. doi:10.26434/chemrxiv-2025-235nn.

[80] M. Sigot and S. Tassoti. A missed opportunity for no-code chatbots? current challenges in publicly available chemistry gpts. *Journal of Chemical Education*, 102 (5):2151–2159, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.4c01552.

[81] L. Sizemore, B. Hutchinson, and E. Borda. Use of machine learning to analyze chemistry card sort tasks. *Chemistry Education Research and Practice*, 25(2):417–437, 2024. doi:10.1039/d2rp00029f.

[82] R. Soong, B. G. Pautler, A. Moser, A. Jenne, D. H. Lysak, A. Adamo, and A. J. Simpson. Case (computer-assisted structure elucidation) study for an undergraduate organic chemistry class. *Journal of Chemical Education*, 97(3):855–860, 2020. ISSN 0021-9584. doi:10.1021/acs.jchemed.9b00498.

[83] B. Sorenson and K. Hanson. Identifying generative artificial intelligence chatbot use on multiple-choice, general chemistry exams using rasch analysis. *Journal of Chemical Education*, 101(8):3216–3223, 2024. ISSN 0021-9584. doi:10.1021/acs.jchemed.4c00165.

[84] S. M. S. Subasinghe, S. G. Gersib, and N. P. Mankad. Large language models (llms) as graphing tools for advanced chemistry education and research. *Journal of Chemical Education*, 102(4):1563–1571, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.4c01498.

[85] J. Sweller. Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2):257–285, 1988. doi:10.1016/0364-0213(88)90023-7.

[86] V. Talanquer. Interview with the chatbot: How does it reason? *Journal of Chemical Education*, 100(8):2821–2824, 2023. ISSN 0021-9584. doi:10.1021/acs.jchemed.3c00472.

[87] S. Tassoti. Assessment of students use of generative ai: Prompting strategies and prompt engineering in chemistry education. *Journal of Chemical Education*, 101 (6):2475–2482, 2024. ISSN 0021-9584. doi:10.1021/acs.jchemed.4c00212.

[88] M. Taylor, Z. Muwaffak, M. R. Penny, B. R. Szulc, S. Brown, A. Merritt, and S. T. Hilton. The rise of the ai scientist: unleashing the potential of chat-gpt powered avatars in virtual reality digital-twin laboratories. *ChemRxiv*, pages 1–11, 2023. ISSN 2573-2293. doi:10.26434/chemrxiv-2023-t4vg7.

[89] E. S. Thrall, S. E. Lee, J. Schrier, and Y. Zhao. Machine learning for functional group identification in vibrational spectroscopy: A pedagogical lab for undergraduate chemistry students. *Journal of Chemical Education*, 98(10):3269–3276, 2021. ISSN 0021-9584. doi:10.1021/acs.jchemed.1c00693.

[90] E. S. Thrall, F. Martinez Lopez, T. J. Egg, S. E. Lee, J. Schrier, and Y. Zhao. Rediscovering the particle-in-a-box: Machine learning regression analysis for hypothesis generation in physical chemistry lab. *Journal of Chemical Education*, 100 (12):4933–4940, 2023. doi:10.1021/acs.jchemed.3c00765.

[91] E. S. Thrall, O. M. Vanden Assem, J. A. Schneider, J. Schrier, and S. Tassoti. Can you help chatgpt get an "a" in organic chemistry? teaching effective prompting of large language models for reaction prediction. *Journal of Chemical Education*, 103 (4):2373–2378, 2026. doi:10.1021/acs.jchemed.5c01712.

[92] M. Thway, J. Recatala-Gomez, F. S. Lim, K. Hippalgaonkar, and L. W. T. Ng. Harnessing genai for higher education: A study of a retrieval augmented generation chatbot's impact on learning. *Journal of Chemical Education*, 102(9):3849–3857, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00113.

[93] R. Tran, V. Brunskill, A. Musgrove, T. C. Sutherland, and D. J. Derksen. A problem-based introduction to machine learning in the undergraduate organic chemistry laboratory: Prediction of diels-alder reaction rates. *Journal of Chemical Education*, 102(8):3443–3451, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00313.

[94] J. Tyson. Shortcomings of chatgpt. *Journal of Chemical Education*, 100(8):3098–3101, 2023. ISSN 0021-9584. doi:10.1021/acs.jchemed.3c00361.

[95] S.-Ş. Uçar, I. Lopez-Gazpio, and J. Lopez-Gazpio. Evaluating and challenging the reasoning capabilities of generative artificial intelligence for technology-assisted chemistry education. *Education and Information Technologies*, 30(8): 11463–11482, 2025. doi:10.1007/s10639-024-13295-6.

[96] B. R. Varju, M. E. Noman, A. P. Dicks, and D. S. Seferos. Incorporating chatgpt into a simulated peer-review assignment in an upper-level materials chemistry course. *Journal of Chemical Education*, 103(1):258–265, 2026. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00790.

[97] V. Venkatasubramanian. Teaching artificial intelligence to chemical engineers: Experience from a 35-year-old course. *Chemical Engineering Education*, 56(4), 2022. doi:10.18260/2-1-370.660-130423.

[98] M. A. Vincent and B. J. Lear. Using exam preparation and reflection to introduce artificial intelligence tools in honors general chemistry. *Journal of Chemical Education*, 102(10):4470–4478, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00618.

[99] M. B. Walsh, C. M. Moss, B. G. Johnson, D. A. Holder, and J. D. Madura. Quantitative of a cognitive modeling intelligent tutoring system on student performance in balancing chemical equations. *Chemical Educator*, 7(6):379–383, 2002. ISSN 1430-4171. doi:10.1007/s00897020620a.

[100] R. Wang. Ai + chemistry education in top chinese university: embedding computational thinking into curriculum. *ChemRxiv*, pages 1–18, 2025. ISSN 2573-2293. doi:10.26434/chemrxiv-2025-8kh8t.

[101] S. Wang, F. Chen, G. Chen, H. Shi, Y. You, J. Guo, S. Zheng, and X. Zhu. Using local llm tools to optimize chinese high school chemistry education: Practice, challenges, and future directions. *Journal of Chemical Education*, 102(10):4368–4375, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00220.

[102] X. Wang, L. Zhang, Y. Mao, D. J. McGillivray, and Z. Wang. Chedu: A guided ai teaching assistant for chemistry education and exam support. *Journal of Chemical Education*, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c00036.

[103] F. M. Watts, A. J. Dood, G. V. Shultz, and J.-M. G. Rodriguez. Comparing student and generative artificial intelligence chatbot responses to organic chemistry writing-to-learn assignments. *Journal of Chemical Education*, 100(10):3806–3817, 2023. ISSN 0021-9584. doi:10.1021/acs.jchemed.3c00664.

[104] X. Xie, S. Ma, Y. Li, C. Shang, and Z. Liu. Application of machine learning potential-based theoretical simulations in undergraduate teaching laboratory course design. *Daxue Huaxue*, 40(3):140–147, 2025. doi:10.12461/pku.dxhx202405164.

[105] M. Xue, B. Liu, and X. Huang. Machine learning in chemistry: a data centred, hands-on introductory machine learning course for undergraduate students. *ChemRxiv*, pages 1–20, 2025. ISSN 2573-2293. doi:10.26434/chemrxiv-2025-9zldf.

[106] E. Yacobson, Y. Schleifer, Z. Bar-Dov, S. Rap, R. Blonder, and G. Alexandron. Benchmarking ai on standard chemistry exams: Llms still underperform compared to high school students. *Journal of Science Education and Technology*, 2026. ISSN 1573-1839. doi:10.1007/s10956-026-10310-y.

[107] B. J. Yik and A. J. Dood. Chatgpt convincingly explains organic chemistry reaction mechanisms slightly inaccurately with high levels of explanation sophistication. *Journal of Chemical Education*, 101(5):1836–1846, 2024. ISSN 0021-9584. doi:10.1021/acs.jchemed.4c00235.

[108] B. Yildirim and A. T. Akcan. Ai-professional development model for chemistry teacher: Artificial intelligence in chemistry education. *Journal of Education in Science, Environment and Health*, pages 161–182, 2024. doi:10.55549/jeseh.741.

[109] J. D. Young, L. Dawood, and S. E. Lewis. Chemistry students' artificial intelligence literacy through their critical reflections of chatbot responses. *Journal of Chemical Education*, 101(6):2466–2474, 2024. doi:10.1021/acs.jchemed.4c00154.

[110] E. Yuriev, D. J. Wink, and T. A. Holme. Virtual special issue call for papers: Investigating the uses and impacts of generative artificial intelligence in chemistry education. *Journal of Chemical Education*, 100(9):3168–3170, 2023. ISSN 0021-9584. doi:10.1021/acs.jchemed.3c00829.

[111] E. Yuriev, D. J. Wink, and T. A. Holme. The dawn of generative artificial intelligence in chemistry education. *Journal of Chemical Education*, 101(8):2957–2959, 2024. ISSN 0021-9584. doi:10.1021/acs.jchemed.4c00836.

[112] E. Yuriev, M. Orgill, and T. Holme. Generative ai in chemistry education: Current progress, pedagogical values, and the challenge of rapid evolution. *Journal of Chemical Education*, 102(9):3773–3776, 2025. ISSN 0021-9584. doi:10.1021/acs.jchemed.5c01105.

[113] J. Zhou. Design and practice of comprehensive computational chemistry experiment based on high-throughput computation and machine learning. *Daxue Huaxue*, 40(9):69–75, 2025. ISSN 1000-8438. doi:10.12461/pku.dxhx202411067.